

DOCUMENT RESUME

ED 209 710

CS 503 665

AUTHOR Myrick, Howard A.; Keegan, Carol
TITLE Boston (WGBH) Field Testing of a Qualitative
 Television Rating System for Public Broadcasting.
INSTITUTION Corporation for Public Broadcasting, Washington, D.C.
 Office of Communication Research.
REPORT NO ISEN-0-89776-071-9
PUB DATE Jun 81
NOTE 59p.

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS Adults; *Evaluation Methods; Field Tests; Predictive
 Validity; *Programing (Broadcast); *Public
 Television; Rating Scales; *Television Research
IDENTIFIERS Media Use

ABSTRACT

A field test was conducted to examine the efficiency of a paper-and-pencil television diary for collecting viewer evaluations of public and commercial television programs. Six hundred individuals in Boston, Massachusetts, kept qualitative viewing diaries for one week. The results indicated high and enthusiastic cooperation by the total sample. The diary was determined to be easy and convenient to use, with acceptably low nonresponse rates on each of the four qualitative scales used to evaluate programs. The resulting profile for public television lent credence to the notion that public television programs cannot and need not be evaluated solely on the basis of audience size. In addition, the data collected proved highly relevant to program development, scheduling, promotion, and evaluation, both for the local Boston station and for stations at the national level. Sample pages from the diary are appended.
 (Author/FL)

 * Reproductions supplied by EDRS are the best that can be made *
 * from the original document. *

ED209710

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

+ This document has been reproduced as
received from the person or organization
originating it.

☐ Minor changes have been made to improve
reproduction quality.

• Points of view or opinions stated in this docu-
ment do not necessarily represent official NIE
position or policy.

Boston (WGBH) Field Testing
of
A Qualitative Television Rating System
for
Public Broadcasting

A Report Prepared
by
Howard A. Myrick, Ph.D.
and
Carol Keegan, Ph.D.

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY
Howard A. Myrick

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

Office of Communication Research
Corporation for Public Broadcasting

June 1981

CS 583665

©Copyright 1981
Corporation for Public Broadcasting
1111 Sixteenth Street N.W.
Washington, D.C. 20036
An Equal Opportunity Employer - M/F
ISBN:0-89776-071-9

TABLE OF CONTENTS

PREFACE.....	iv
ACKNOWLEDGEMENTS.....	v
EXECUTIVE SUMMARY.....	1
I. INTRODUCTION.....	2
II. APPLICATIONS AT THE LOCAL STATION LEVEL.....	9
III. IMPLICATIONS FROM A NATIONAL PERSPECTIVE.....	16
IV. METHODOLOGICAL ASSESSMENT.....	31
V. CONCLUSIONS.....	44
APPENDIX A: QUALITATIVE DIARY EXCERPTS	
APPENDIX B: BIBLIOGRAPHY	

PREFACE

The Corporation's active role in the area of qualitative television research has been supported throughout by the enthusiastic curiosities of public television programming executives and policymakers who possess the kind of philosophical vision that originally created public television.

Rebounding from the painful notion of change, and endorsing the moral obligations of public service accountability, these industry producers, programmers, policymakers and researchers have helped motivate the Corporation's search for comprehensive research measures tailored to the programming missions of public television, and to the articulate sensibilities of its viewership.

Moved by such enthusiasm, we brought much responsibility and ambition to the pilot research described in this document. Acknowledging that no one field test of a qualitative television diary can raise or resolve every methodological issue, we do hope this document serves to sustain the industry's search for more complete knowledge of its past programming achievements and its future programming opportunities.

Howard A. Myrick, Ph.D.
Director
Office of Communication Research

Carol Keegan, Ph.D.
Associate Director
Office of Communication Research

ACKNOWLEDGEMENTS

A grateful acknowledgement must be made of the research expertise so skillfully contributed by Dr. Joseph Philport of the Arbitron Company. His own research in qualitative ratings and his ability to adapt research procedures to accommodate special information needs in public television were important contributions to this effort.

Sincere gratitude is also extended to the Community Advisory Board of WGBH and its station coordinator, Beth Pile. Without the clarity of this body's approach to programming philosophies and responsibilities in public television, and without their enthusiastic receptivity to innovative audience research concepts, this field test could never have been accomplished.

Finally, very special thanks are extended to Dr. D. Liroff, Broadcast Manager at WGBH, for his sustained confidence and support throughout the conduct of this somewhat risky venture, for the honesty and sense of public service with which he approached these new accountability measures, and for his intellectual agility in handling both methodological and applications issues.

EXECUTIVE SUMMARY

A field test was conducted, to examine the efficiency of a paper-and-pencil television diary for collecting viewer evaluations of public and commercial television programs. Six hundred individuals in Boston kept qualitative diaries for one week in November, 1980. In cooperation with WGBH, the public television station in Boston, the Office of Communication Research designed the diary, the sample, and data collection and analysis methods to maximize the relevance of resulting qualitative data to a variety of local and national programming activities.

Field test results indicated high and enthusiastic cooperation by the total sample. The diary itself proved easy and convenient for the sample to keep, with acceptably low nonresponse rates on each of the four qualitative scales used to evaluate viewed programs. Resulting qualitative profiles for public television lent empirical credence to the notion that PTV programs cannot and need not be evaluated solely on the basis of audience size. Furthermore, the qualitative data collected proved highly relevant to program development, scheduling, promotion and evaluation activities at the local station and at the national level.

The report identifies several research design modifications which will further improve the efficiency of the data collection procedures and enhance the applied value of resulting data for programmers and policymakers in public television.

I. INTRODUCTION

A. Overall Organization of the Field Test

With the field test described in this document, the Office of Communication Research (OCR) culminates a three-year investigation of the potential application of qualitative television ratings to public broadcasting activities and objectives. Earlier reports by OCR in this area (CPB, 1980a, 1980b) had reviewed scholarly and industrial qualitative research initiatives here and abroad, and found an impressive range of potential and actual applications of qualitative ratings data. The field test described herein attempted to gauge the applied value of such data to a variety of public television (PTV) programming activities and goals, at both local and national levels.

Accordingly, the report is divided into three parts which separately address major areas of concern often raised whenever the potential value of qualitative data is discussed. The first section of this report by David Liroff, Broadcast Manager at WGBH, assesses the value of the data to a local PTV station which must make daily decisions about program development, scheduling, promotion and station accountability to the community. The second part of the report scrutinizes the same qualitative data to study their worth in national concerns related to program development and funding, institutional commitment to minority and special interest audiences, and overall system accountability. Finally, in the third part of this report, Joseph Philport, project manager at Arbitron, evaluates the methodological integrity of the research procedures employed in the field test, concerning himself with the development of a cost-effective and valid methodology for collecting and analyzing viewer opinions.

Since the report was originally conceived and designed with these three levels of evaluation in mind, the research methodology cast a rather wide net. It was designed to reveal as much as possible about valid and efficient procedures for studying viewer reactions to television programs, and practical applications of resulting data to both local and national programming responsibilities in public television. While practical research methods and "real-world" applications heavily influenced the research design, the investigation was primarily driven by those programming philosophies, mandates and objectives uniquely associated with public broadcasting in this country. We will briefly review the reasoning which originally convinced OCR of the potential relevance of qualitative research to special information needs in public television. More extensive discussions of the philosophical heritage of these innovative research methods is provided in earlier OCR publications (CPB, 1980a, 1980b).

B. The Need for Innovative Audience Research Methodologies in Public Television

Because public broadcasting is driven by special mandates and programming objectives ("excellence," "innovation," "special interest" and minority audience service, "diversity," "alternative" programming), it has always been insufficient to assess the value and effectiveness of public television programs with estimates of the size and composition of PTV audiences. Traditional ratings systems, developed principally to serve the information needs of the commercial broadcaster, have left PTV programming decision-makers almost completely uninformed of their overall programming effectiveness and of their future program and audience development options. That is, traditional ratings data have proven inadequate measures of public television's accountability to the public it was created to serve, and they have provided broadcasters and producers with insufficient information on how best to develop, distribute, schedule and promote public television programs.

The unsatisfied information needs of the public broadcaster are as complex as his or her programming objectives. They parallel the programming life cycle and aim at: generating new programming ideas and directions, building the public television audience, improving program production, and making balanced and accountable decisions about the effectiveness of individual programs.

Traditional measures of audience size and demographics tell us little or nothing about the extent to which PTV programs serve and satisfy their viewers, about "gaps" in current public and commercial television content, or about "alternative" services public television might provide because they would otherwise go unaddressed by other sources of programming.

Traditional program ratings tell us little or nothing about how to satisfy the viewing needs and interests of minority and special interest audiences too small to be targeted by commercial television, or about means for fostering diversity in PTV program content, or about the viewing preferences of minority and special interest audiences.

Ratings information limited to audience size and composition tell us little or nothing about program concepts, treatments or formats likely to attract new viewers, or about desirable aspects of PTV programs that can be promoted to counteract undesirable PTV images among nonviewers, or about programming directions likely to attract new audiences without alienating the current PTV viewership.

Traditional program ratings tell us little or nothing about the effectiveness of PTV in serving special interest audiences which cannot be demographically characterized, or about how to use viewing behavior and attitudes of special interest audiences to cross-promote targeted programs, avoid detrimental counter-programming, and maximize audience flow.

Traditional program ratings also fail to provide comprehensive evaluation criteria for deciding to air and/or fund a given program, to assess whether a station is accountable to its community, or whether the entire system is fulfilling its objectives.

Compressing all of these research motivations into a general purpose, OCR determined that the new system to be tested should maximize our knowledge of audience reactions to current programs, and identify specific program development, scheduling and promotion activities most likely to enhance viewer satisfaction with future PTV programming.

C. Development of the Methodology

Through a highly fortunate confluence of interests, OCR found WGBH's Community Advisory Board (CAB) grappling with similar issues of PTV accountability criteria and the role of audience research in enhancing station service to Boston's community.

Early in September 1980, OCR met with the CAB's research task force to suggest a collaborative research effort, in which both national and local utilities of a qualitative ratings system could be field tested. In cooperation with David Liroff, Broadcast Manager at WGBH, Beth Pile, CAB Coordinator at WGBH, and the CAB's research task force, OCR began to design an experimental qualitative television diary field test.

Primarily, the system was designed to surpass traditional demographic audience descriptions and encompass the special programming objectives which distinguish public television from its commercial counterparts. In addition to evaluating current PTV programs, it was determined the system should be forward-looking, identifying new and promising program content areas. Philosophically, the research design tried to strike a middle ground between the paucity of traditional viewing data and the detailed formative and summative information usually addressed in special studies. The waking appetite of programming and policymaking executives for information which went "beyond" traditional ratings data produced a voracious appetite for comprehensive audience information, and a reversal of traditional research deficiencies in one swift study. The middle ground, where we would learn more--but not everything-- about our programming and our audiences, was the territory OCR chose to cover.

Beyond such information needs, several methodological concerns heavily influenced the research design. Three deficiencies of traditional ratings systems inspired remedial features in the new system. First, it was determined that the new system should oversample minorities, to allow sufficient sample size for programs targeted to black and Hispanic audiences, and to better represent minority group reactions to general audience programs. Second, if

representative opinions of PTV viewers were to be collected in regard to individual PTV programs, then the new system should oversample PTV viewers to increase the likely sample size for specific PTV programs. Finally, since many PTV programs and programming directions involve target audiences which can be characterized psychographically (e.g., by interest in subject matter) but not necessarily demographically, the new system should characterize viewer subgroups by their degree of interest in subject matter treated in current PTV programs or potentially treated in future programs.

Under contract with the Arbitron Company, specific details of the research design were finalized and the field test was conducted during the week of November 19-25, 1980.

Instrument

One of the first considerations was the method of data collection. Cost factors, the need for acceptable cooperation rates and unobtrusive procedures, and the proven experiences of PTV research colleagues in other countries all combined to make paper-and-pencil diaries the preferred data collection instrument for the pilot test. The traditional Arbitron diary format was modified, to add four qualitative rating scales to records of half-hour viewing behavior. The four qualitative rating scales selected ("entertaining," "different from most television programs," "useful to me," "informative") were chosen both to tap important viewer evaluation dimensions identified in more basic research studies (Landis, 1979; Himmelweit, 1980) and to reflect the PTV industry's special programming objectives. Instrument instructions encouraged respondents to view as they normally do, and no PTV sponsorship of the study was indicated. The diary was constructed in half-hour units covering 3-11PM for each day of the survey week.

Viewers recorded the name and source (channel number and station call letters) for programs they decided to view between 3 and 11 PM during the survey week, and gave four qualitative ratings for each program viewed.

Additional sections of the diary covered respondent motivations for viewing television, degree of interest in 22 special interest areas, and demographic characteristics.

Possible motivations for viewing television (Appendix A, p. A-1) were selected to study larger orientations toward television which might differentiate PTV viewers and nonviewers, suggest why nonviewers avoid PTV, and indicate how promotional activities and production qualities might help overcome such barriers to viewing. Motivations covered in the instrument were:

- to learn about people, places and things
- to keep up with current issues and events

to be exposed to the arts
to give me things to talk about with other people
to be entertained
to relax and pass the time

The 22 special interest areas studied in the diary were selected to attempt a psychographic segmentation of special interest audiences. If viewers and nonviewers could be categorized in this manner, current programs could be evaluated in terms of their ability to satisfy these special interest audiences. Furthermore, future program concepts and directions could be proposed in cases where special interest audiences were either not served or not satisfied with current CTV or PTV programs.

Sample

The sampling frame consisted of persons 12 years of age and older living in households which had returned 11-tab diaries from recent Arbitron Television Market Report surveys in the Boston market. A random interval selection technique was modified to draw samples with a sufficient number of PTV viewers (of individual PTV programs and of PTV generally). The sampling plan was designed to reach a quota of 350 PTV households (those who, according to prior Arbitron diary periods, had viewed a PTV station for at least two quarter hours a week) and 150 nonviewers (households in which prior Arbitron diary periods indicated no PTV viewing in the week surveyed).

Special oversampling of minorities, to permit comparative analyses of their viewing habits and preferences, set a minimum quota of 50 black and 50 Hispanic respondents.

Procedures

Diary placement telephone calls were made by Arbitron interviewers to each sample household being recruited. The interview established whether the number of eligible persons in the household justified inclusion in the total sample, given the special quota sampling of PTV viewers and nonviewers, and of a number of racial and ethnic viewer categories.

If the household met sampling requirements and volunteered cooperation, a diary was mailed to each resident twelve years of age or older, along with a \$1 incentive. To encourage subsequent cooperation during the week, a reminder post card was mailed to each cooperating household, urging them to complete and return their diaries.

Data Analysis and Reporting

Before proceeding with the three analyses of the data which follow, (local utility, national utility, methodological integrity), the data reporting practices used in this field test must be explained.

Viewers evaluated four qualitative aspects of programs they viewed along the following four-point scale:

1	2	3	4
"not at all"	"somewhat"	"very"	"extremely"

Across all programs evaluated by the panel, average qualitative scores were calculated for each dimension:

entertaining	2.697
different	2.444
useful	2.018
informative	2.220

With these mean scores, individual program ratings along the four scales could be indexed, to show how they compare with "average" qualitative ratings for all programs rated by the panel. This indexing procedure means that each of the qualitative ratings for a program is divided by the mean rating for all programs evaluated by the panel along the same scale. As a result, an index of 100 means the program merited a rating equal to the average panel rating for all programs. Any index less than 100 means the program was rated below the average qualitative rating for all programs, and the smaller the number, the more the program rating is below average. Conversely, the more a program rating exceeds 100, the more its viewers gave it an above average rating.

The same indexing method was used to examine panel viewing levels for all programs covered in the diary. The viewing index for any program equals the number of panel members who watched it, divided by the average number of panel viewers* for all programs covered in the study.

While the analysis of qualitative ratings data can involve far more complex statistical manipulations (see Section IV), the pilot test analyses were targeted for non-researchers and selected to enhance understanding and the applied value of the data.

*Because amount of viewing varied considerably across different demographic categories, viewing indices for individual programs were based on an average audience size calculated separately for each demographic subgroup, as well as for the total sample.

In Section II, David Liroff, Broadcast Manager at WGBH, who helped OCR tailor this qualitative field test to his station's special information requirements, evaluates the applied value of qualitative television ratings to the operational and policy concerns of a local PTV station. Specifically, Liroff was asked whether the availability of these new audience data might in any way enlighten a station's programming decisions. His responses illustrate how new qualitative audience information might provide critical guidance in a variety of programming decision areas.

II. APPLICATIONS AT THE LOCAL STATION LEVEL

- David Liroff, Ph.D., WGBH Broadcast Manager

Public television program directors frequently make assumptions about their audiences based on sparse evidence. Limited quantitative data from syndicated research and qualitative feedback gleaned from a few letters, phone calls and conversations often provide the only empirical data on which to base decisions about program acquisitions, production, scheduling and promotion. As our colleagues in commercial broadcasting continually demonstrate, there is no scientific approach guaranteed to attract audiences to programs, although for them the general objective of maximizing the number of viewers for each program in their schedules is relatively easy to define. The public broadcaster, on the other hand, is charged with achieving numerous, often conflicting objectives which have been prescribed variously by legislators, critics, funders, and by virtually any constituency which has looked to public television to provide that which commercial television has not. Among the mandates are: the provision of quality, alternative, and innovative programs to special interest audiences which may be small; granting access to individuals and groups whose viewpoints and works have received little exposure in the mass media; and making formal and informal educational opportunities more widely available, while simultaneously increasing the number of viewers served by each station and, not incidentally, surviving financially.

Assuming that these mandates and others are assigned appropriately to public television, how is progress toward their achievement to be measured? In the absence of appropriate measurable performance standards, the PTV program director is virtually defenseless when confronted by determined partisan advocates for special interests who claim, at times appropriately, that their particular special needs are not being met by the station. The absence of these standards also makes it difficult for the programmer to evaluate efforts to extend services to new audiences, or to identify aspects of station programming which should be improved or deleted.

While it is clear that not all of the mandates lend themselves to neatly quantifiable performance standards, the pilot test reported in this document does allow special interest audience groups to evaluate programs across several qualitative scales, and--where differences among various segments of the audience are important to analyze--it does permit a closer examination of specific target audience response.

The following are examples of the types of potential applications of these data. It must be stressed that due to the small size of the sample in this pilot study, any conclusions about the performance of the specific programs must be considered extremely tentative. Data are reported for illustrative purposes only.

Example I.

The second season of This Old House, a series on home renovation, featured the conversion of a run-down estate into five condominium units. At the time the series was broadcast, condominium conversion was a controversial issue in the Greater Boston area. An extreme shortage of low and moderate income rental housing was being exacerbated by the conversion of older apartment buildings into high-priced condominiums. A few critics charged that the television series appealed to high income viewers and ignored the needs and interests of low income inner city residents threatened with displacement.

What audience segments perceived
the highest degree of quality
in This Old House?

	<u>Entertainment</u>	<u>Different</u>	<u>Useful</u>	<u>Informative</u>
General Audience	103	147	179	166
Upper Income	93	136	173	158
Middle Income	102	139	168	162
Lower Income	<u>111</u>	<u>152</u>	<u>188</u>	<u>171</u>
Blue Collar	<u>111</u>	<u>164</u>	<u>198</u>	<u>180</u>
Professional	93	143	173	158

Lower income and blue collar viewers gave the program the highest ratings across the four qualitative scales. The critics did not accurately represent their feelings.

Viewing levels varied across the three income groups as well, with somewhat heavier penetration among lower income viewers:

	<u>% Viewing</u>
Lower Income	3.4%
Middle Income	2.6%
Upper Income	2.1%

Example II.

Say Brother is a locally produced weekly series targeted at Boston's Third World Community, the largest component of which is black. Nielsen surveys report small audiences for the series, in part the result of averaging viewing across all households, the vast majority of which are not in the target audience.

Does the target audience for
Say Brother value it more highly
than other audience segments do?

	<u>Entertainment</u>	<u>Different</u>	<u>Useful</u>	<u>Informative</u>	<u>% Viewing</u>
PTV Viewers	74	106	109	99	1.2%
Blacks	74	102	<u>149</u>	<u>135</u>	<u>3.8%</u>
Upper Income	74	109	<u>83</u>	<u>75</u>	<u>1.6%</u>
Managers	37	41	50	45	1.7%
Race Relations/Ethnicity (audience interest segment)	<u>111</u>	<u>123</u>	99	90	0.9%

Clearly, the target audience and a related self-designated audience interest segment evaluate the program differently than a sampling of other audience segments. The qualitative evaluations provide added dimension to an understanding of the program's performance.

Example III.

The Ten O'Clock News is broadcast each weeknight on WGBH-TV. Compared with other local newscasts in the market, the program places much greater emphasis on national and international news, and provides fewer but more detailed reports about local issues.

How do those interested in National and
International News/Politics rate the information
quality of the The Ten O'Clock News and other news/public
affairs programs?

CBS Evening News	151
Washington Week in Review	149
60 Minutes	145
NewsCenter 5 Late Edition	141
Ten O'Clock News	127
Bill Moyers' Journal	113

Again, a warning about small sample size. In this case, the Ten O'Clock News ratings were averaged over the five respondents in this interest segment who had viewed the program, and we cannot assume they are representative of the entire interest segment. However, if the data were reliable, and if evaluations by other audience segments across this and other qualitative scales rated this program below average in comparison with other news and public affairs programs, it might indicate that the program was not achieving its objectives. Actions to be taken might include a reassessment of the program's content and

format, consideration of alternative approaches to the presentation of similar types of information, or a reallocation of resources to other programs. Additional audience research focusing on this particular series might reveal other factors contributing to the viewers' assessment of its quality.

Example IV.

Various audience segments are often mistakenly assumed to be monolithic in their television programming tastes. In fact, the complex motivations and behaviors of individual viewers will not be found uniformly distributed across all members of a demographic or interest segment. For example:

What PTV programs are evaluated most highly by blue collar viewers?

Entertaining

Wild, Wild World of Animals	131
The Men Who Made the Movies	124
Cosmos	111
This Old House	111
Sneak Previews	111

Different

This Old House	164
Masterpiece Theatre	164
Nova	123
Survival	123
Wild, Wild World of Animals	123

Useful

This Old House	198
Nova	173
Masterpiece Theatre	149
Cosmos	149
Sneak Previews	149

Informative

This Old House	180
Masterpiece Theatre	180
Cosmos	180
Nova	158
Wild, Wild World of Animals	146

There is a broad mix of program content contained in these lists, defying pat stereotypical description of the program preferences of the viewers in this audience segment.

Example V.

There is a common assumption among television professionals that viewers will evaluate elaborate productions more favorably than those with simpler production values, (thus the prejudice against "talking head" programs.) The Tom Cottle Show on WGBH is a half-hour interview program dealing with the problems which individuals and families confront in everyday life. A thematically similar program on commercial television is The Baxters, whose format includes a fifteen minute drama followed by a highly edited studio audience discussion.

How do viewer evaluations of WGBH special interest programs compare with evaluations of special interest programs on commercial television?

Adult Audience

	<u>Entertaining</u>	<u>Different</u>	<u>Useful</u>	<u>Informative</u>
<u>The Tom Cottle Show</u>	102	123	136	146
<u>The Baxters</u>	95	91	99	94

Interest Segment: Concern for Persons with Physical or Emotional Disabilities

<u>The Tom Cottle Show</u>	111	143	173	180
<u>The Baxters</u>	95	82	113	103

The production complexity of the two programs does not appear to be a critical factor in viewer evaluations. Note too that within the target interest segment, qualitative reactions accentuate differences perceived between the two programs by the total adult audience.

Conclusion

Although an analysis of the data does not demonstrate that membership in a demographic or social interest group is a necessary and sufficient predictor of viewer behavior or content evaluation, it appears that the qualitative ratings methodology tested in the pilot study has potential as an instrument to evaluate the effectiveness of a station's service to particular special interest audiences, using qualitative as well as quantitative criteria.

The data produced are not sufficient to support a detailed analysis of individual programs, but the process does generally identify strengths and weaknesses in the station's overall delivery of services to audiences. Additional research (through the use of focus group discussion, or recruited viewer panels, for example) can then be employed to more fully diagnose viewer attitudes about specific programs.

Clearly, the implementation of a qualitative ratings system would substantially increase the program director's ability to evaluate the effectiveness of a station's service by providing detailed audience information not previously available. As the pilot study demonstrates, performance standards appropriate to some of the objectives of public television can be developed.

The next section of this report assesses the extent to which data of the type collected in the Boston field test can satisfy the national information needs which originally inspired OCR's qualitative research program.

As in the Liroff paper which precedes it, this section explores applications of qualitative data and refrains from addressing basic research questions needed to establish the validity and reliability of the field test data. Such methodological issues are treated in Section IV.

III. IMPLICATIONS FROM A NATIONAL PERSPECTIVE

While, in the preceding section, David Liroff was able to pose direct questions about the applicability of the Boston pilot data to local station information needs, national questions of system-wide performance obviously can only be examined in a hypothetical manner (i.e., assuming additional markets and/or a national panel of qualitative diary keepers). Accordingly, throughout Section III, the Boston data are studied as illustrations of the kinds of analyses and interpretations possible, if qualitative data were collected on a national basis. No literal conclusions about national evaluations of PTV programming can be drawn in the pages to follow. However, given the field test demonstration of cost-effective data collection methods for qualitative research (see Section IV), it is instructive to test data relevance to policy and operational questions at the national level.

Question I.

Assuming the field test data were available from a national sample of PTV viewers and nonviewers, what could we learn about the extent to which PTV programs fulfill industry mandates (e.g., education, providing alternative service, fostering diversity and innovation)?

While qualitative data of the type treated in this field test can never provide complete answers to system performance questions, they can generate new and highly relevant information from the viewer's perspective. Overall, just how do PTV viewers evaluate PTV programs? How do these evaluations compare with qualitative ratings for CTV programs?

Figure 1 shows the results of a simple tabulation of the number of PTV and CTV programs whose qualitative ratings, when indexed against average qualitative ratings for all programs evaluated by the total sample, are above average. The reader will recall from Section I that the average scores across all programs evaluated ranged between points "2" and "3" on the four-point scales:

Entertaining	2.697
Different	2.444
Useful	2.018
Informative	2.220

Examining Figure 1, we can conclude that a similar proportion of PTV and CTV programs are perceived as above average in entertainment; however, considerably more PTV programs are seen as different, informative and

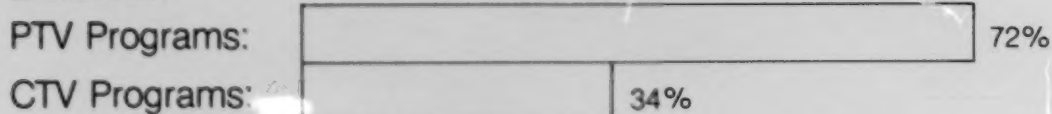
Figure 1.

How do the audiences of PTV and CTV programs feel about the quality of the programs they view?

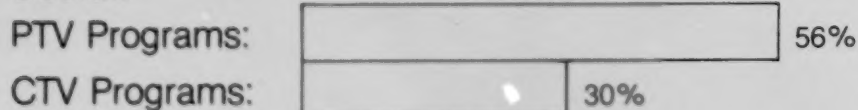
Entertaining:



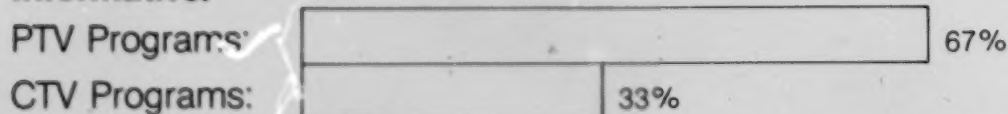
Different:



Useful:



Informative:



Note: Each percent indicates the percent of PTV or CTV programs that were perceived to be above average in a respective qualitative dimension by viewers of the programs.

useful.* Twenty-four out of 43 PTV programs (55.8%) were rated above the norm for all programs in entertainment value, compared to 31 out of 52 CTV programs (59.6%). Thirty-one out of 43 PTV programs (72.1%) were rated above the norm for all programs in terms of being "different from most TV programs," compared to 18 out of 52 CTV programs (34.6%). Twenty-four out of 43 PTV programs (55.8%) were rated above the norm for all programs in terms of being "useful" to the viewer, compared to 14 out of 52 CTV programs (26.9%). Twenty-nine out of 43 PTV programs (67.4%) were rated as above the norm for all programs in terms of being "informative," compared to 17 out of 52 CTV programs (32.7%).

It is also useful to ask what programs, commercial and public, received the highest qualitative ratings on each of the four qualitative indices. "Top Ten" programs** for each qualitative index and for viewing (indexed against the total sample average of 38 viewers per program) are listed on the following page.

* Calculation of means and standard deviations would have revealed that the tabulations in Figure 1 mask superior PTV program performance, with above average PTV program indices significantly higher than those for above average CTV programs. However, the small sample size for many individual programs precluded any meaningful use of more sophisticated data analyses. These calculations await the development of larger PTV samples for individual programs, perhaps by aggregating individual market data for a program across markets in a national panel, or by developing a "qualitative cume" for a program (when possible) over several diary sweeps.

** Omitting any program with a viewing index of less than .10 (4 viewers). Again, sample size for natural viewing of PTV programs, in this one market and with the sampling procedures employed, prohibits meaningful data analysis for many PTV programs.

<u>Entertaining</u>		<u>Different</u>		<u>Useful</u>		<u>Informative</u>		<u>Viewing</u>	
1. Miller's Court ^{cl}	122	This Old House	147	This Old House	179	Old House	166	Dallas ^c	433*
2. Survival	122	Miller's Court ^{cl}	146	Body Works ^{cl}	159	Body Works ^{cl}	154	60 Minutes ^c	297
3. Muppet Show ^c	122	Muppet Show ^c	139	Miller's Court ^{cl}	156	Survival	153	Eight Is ^c Enough	223
4. Connections	117	Survival	137	Sesame Street	155	Nova	153	Love Boat ^c	207
5. Wild, Wild World of Animals	117	Connections	132	Fight Back ^c	153	Wall Street Week	152	Little House ^c On the Prairie	170
6. Quincy ^c	116	Sesame Street	131	Connections	152	Great Performances	150	Taxi ^c	170
7. Evening at the Symphony	116	Cosmos	131	Sneak Previews	146	Sneak Previews	148	Archie's ^c Place	170
8. MASH ^c	115	Nova	126	Nova	141	Fight Back ^c	148	Lawrence Welk ^c	154
9. Dallas ^c	114	Sneak Previews	125	Black News ^{cl}	138	CBS Evening ^c News	148	One Day At ^c A Time	154
10. Masterpiece Theater	114	Masterpiece Theater	124	Tom Cottle ^l	136	Tom Cottle ^l	146	MASH ^c	154**

c = commercial program
l = local program

* The Dallas episode broadcast that week was the intensively promoted program in which J.R. Ewing was shot by an unidentified assailant.

** The highest viewing for a PTV program was only 101 -- Survival.

While no program in the study scored very high on entertainment value, both CTV and PTV programs are scattered throughout the "top ten" entertainment list. However, on the remaining three qualitative scales, we see a predominance of PTV programs, both local and commercial, but only six commercial programs appear anywhere among the top ten programs on the four qualitative scales. Notice that, of the ten programs with the largest audiences, only Dallas and MASH were included among the top qualitative programs.

With such findings, we can contribute important information about system performance in terms of PTV programming missions and philosophies. Given the predominance of PTV programs along the dimensions of difference, informativeness and usefulness, we can argue that alternative service is perceived by the PTV viewer, that PTV is seen as offering innovative and diverse programming (high "different" ratings*), and that it serves educational functions ("informative" and "useful").

Finally, with the special interest group segmentation capability tested in the Boston study, questions of diversity might be approached by asking not only how demographically diverse audiences are, but also how diverse PTV audiences are in terms of the special interests the viewership represents. This new capability is further explored in the response to Question 2 below.

Question II.

How well is PTV currently serving special interest audiences? In terms of future program funding decisions, what new audiences could PTV begin to serve which are not currently served by either CTV or PTV programs?

As explained in Section I above, one unique methodological feature of the pilot test was derived from the PTV mandate to serve diverse special interest audiences whose viewing preferences go unaddressed because they do not merit mass and/or commercial programming attention. In this project, diary panelists expressed their degree of interest in 22 content areas which either are or might be treated in PTV and CTV programming. (See page A-1, Appendix A.) A panelist was said to belong to a special interest group if the respondent claimed to be "very interested" or "extremely interested" in a given content area. A comparison of total sample versus special interest group evaluations of Bill Moyers' Journal illustrates the potential utility of these new data in judging a program's effectiveness. (See Table 1.)

*With increasing penetration of new technologies, it will be necessary for qualitative ratings to assess the offerings of new programming sources, to determine whether PTV programming directions are still perceived as unique, and to suggest (through the special interest segmentation attempted in the field test) what alternative programming functions will remain unaddressed unless PTV provides related services.

TABLE 1

ANALYSIS OF A SPECIAL INTEREST GROUP'S RESPONSE
TO A PROGRAM WITH SPECIAL INTEREST APPEAL

		<u>Bill Moyers' Journal</u>			
	<u>ENTERTAINING</u>	<u>DIFFERENT</u>	<u>USEFUL</u>	<u>INFORMATIVE</u>	
People Interested In News/Politics	124	143	165	173	
Total Adult Sample	90	99	99	116	

Summative applications intended for these data focused on examinations of the diversity of PTV audiences, not just in traditional demographic terms, but also in terms of the psychographic breadth of audiences--that is, the number of special interest audiences who view, and are satisfied by PTV programs. Further summative analyses would determine whether, and, if so, which special interest groups view and prefer PTV to CTV program offerings; identification of special interest groups which showed large discrepancies between their level of viewing and satisfaction with PTV vs. CTV programs would speak to the industry's relative success in providing alternative services. As previous research has argued (Frank and Greenberg, 1980), it is most appropriate for PTV program development and evaluation activities to characterize current and potential viewers in terms of their viewing preferences and interests. Since psychographic diversity is a far more complex and refined characterization of television viewership than demographic information can ever provide, and since PTV program content and program types are far more diverse than commercial programs both within broadcast day parts and for the total broadcast day (OCR, 1979), it is likely that such summative analyses would provide more appropriate and accountable measurements of industry service to diverse audiences.

For formative purposes, the applications concept originally focused on trying to identify special interest audiences which, in terms of their viewing and/or their qualitative responses to available PTV and CTV programs, remain unserved. National planning and program development activities, and national research conducted to assist local, regional and private program development efforts could then identify specific "gaps" in the PTV schedule and help prioritize programming directions likely to fill them. However, the low levels of natural PTV viewing obtained in this one-week study--even with its oversampling of PTV viewers--disallow any definitive investigation of the power of these data in responding to formative and summative issues. (See Section IV for further methodological criticism of this research feature.)

Question III.

With qualitative data from a national sample, what could we say about the success or failure of PTV programs in terms of their ability to serve minority audiences?

When WGBH's Community Advisory Board posed this question from a local perspective, panel data were analyzed to test two operational definitions of performance standards for minority targeted programs.

The first definition derived from a research philosophy that argues against applying general audience evaluations to assessments of minority program performance. With the field test oversampling of minority viewers and with the identification of a special interest group of persons interested in "race relations/ethnicity," it becomes possible to compare viewer reactions to a targeted program within the target group, as well as for the general audience.

The top half of Figure 2 dramatically illustrates how the availability of target group information can provide a very different perspective in the evaluation of the black audience targeted program, Say Brother. As Liroff noted in Section II, examination of total audience data would suggest almost average qualitative performance and a very small audience, but data from panelists show qualitative levels and viewing levels to be considerably higher among persons for whom the program is principally intended.

This differential program response in total audience vs. black audience can also emerge from data on nontargeted programs. The Hitchcock program, while intended for a general audience, apparently serves to attract and satisfy black viewers with noticeably superior performance. Data such as these allow an empirical approach to the knotty question of how to serve minority audiences without necessarily alienating the larger audience.

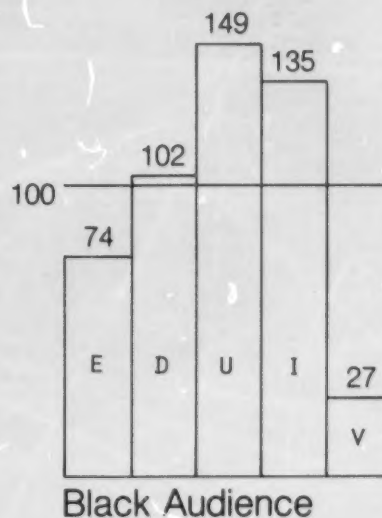
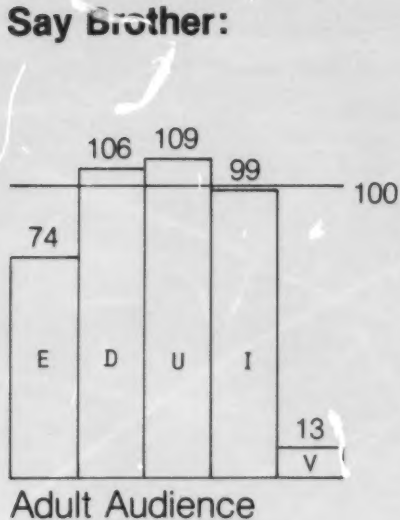
On a national basis, such data could be aggregated across markets to assess the total performance of both targeted and general audience programs in attracting and satisfying minority audiences. Since the field test methodology also permits comparisons with CTV minority and general audience programs, questions of system accountability can be answered with comparative statements about commercial service to minority viewers. Figure 3 demonstrates the kind of local market comparisons which, when aggregated to a national level, can be used to assess PTV vs. CTV minority audience programs.

Figure 2.

E = ENTERTAINING
D = DIFFERENT
U = USEFUL
I = INFORMATIVE
V = VIEWING

Adult Audience vs. Black Audience

Say Brother:



Hitchcock:

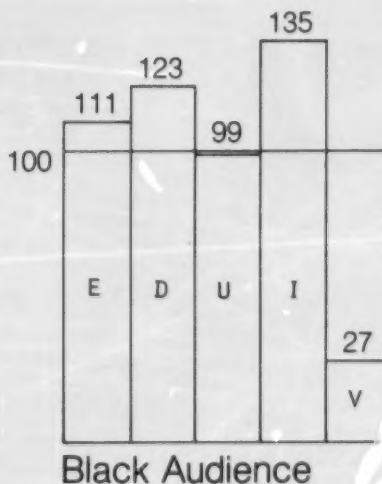
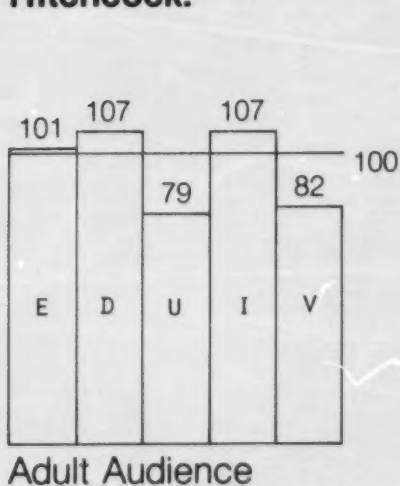
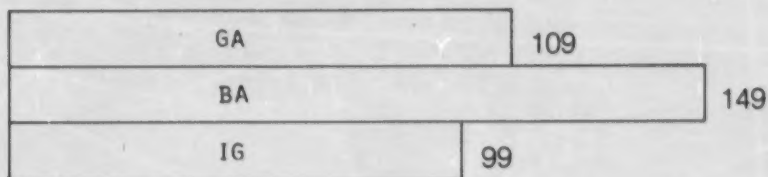


Figure 3.

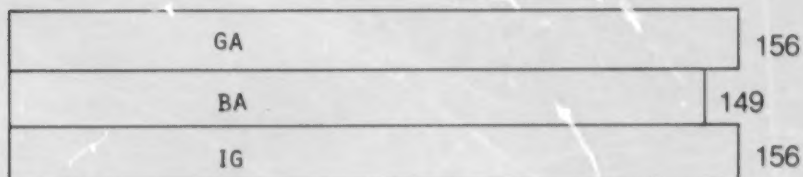
GA = GENERAL AUDIENCES
BA = BLACK AUDIENCES
IG = RACE RELATIONS
INTEREST GROUP

Useful Index

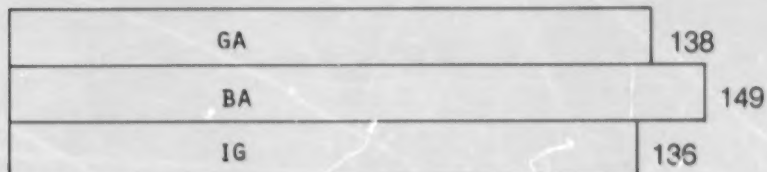
Say Brother



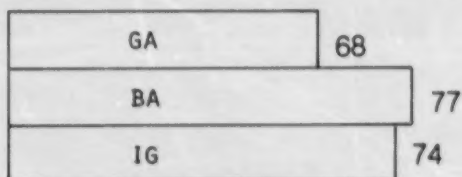
Villa Alegre



Black News
(CTV)



The Jeffersons
(CTV)



Question IV.

PTV is mandated to serve "all the people." Yet the industry is frequently accused of being elitist and aimed at upscale audiences. Are viewers with a lower socio-economic status (SES) less satisfied with PTV than upscale viewers? If so, what programs have downscale appeal, and what program qualities could be emphasized in promotion activities to help broaden the demographic composition of the PTV audience?

Should such formative research questions be studied with the types of qualitative data collected in Boston, a variety of analyses might be conducted to help identify program production, scheduling and promotion activities most likely to attract and/or increase amount of current viewing by lower SES viewers.

Across all PTV and CTV programs, program viewing and evaluation levels could be compared for the following sets of viewers:

blue collar	vs.	professional
less educated	vs.	college educated
low income	vs.	high income

Table 2, for example, compares "top ten" PTV programs by amount of viewing for professional vs. blue collar viewers. First, notice that the PTV viewing indices for blue collar viewers are less than half those of professional viewers, even though panel members from both groups tended to watch equal amounts of television in general. Also notice that three programs which appear exclusively in the top ten list for professional viewers are displaced by the broad-appeal program, Old House, by a multi-ethnic teen targeted program, From Jumpstreet, and by a black-targeted program, Say Brother.

In addition to more obvious program format and content differences implicit in these viewing differences, it is interesting to note qualitative differences in the two group's perceptions of the programs they view in common:

<u>Program</u>	<u>Blue Collar</u>	<u>Professional</u>
Wild, Wild World of Animals	Entertaining: 131	Entertaining: 108
Nova	Useful: 173	Useful: 136
Hitchcock	Informative: 135	Informative: 108
Survival	Useful: 50	Useful: 93

Table 2.

Comparison of Top Ten PTV Programs by Viewing
Indices for Professional and Blue Collar Diary Samples

Blue Collar		Professional	
<u>Program</u>	<u>Index</u>	<u>Program</u>	<u>Index</u>
1. Hitchcock	41	1. Survival	107
2. Wild, Wild World of Animals	36	2. Cosmos	94
3. Matinee at the Bijou	27	3. Hitchcock	67
4. Sneak Previews	27	4. Wall Street Week	67
5. Cosmos	27	5. Washington Week in Review	54
6. Nova	27	6. Masterpiece Theatre	54
7. Survival	14	7. Connections	54
8. This Old House	14	8. Nova	54
9. Say Brother	14	9. Sneak Previews	40
10. From Jumpstreet	14	10. Wild, Wild World of Animals	32
Average number of viewers for all programs	7.29		7.44

Further analyses, not possible with the current data (due to small sample sizes when demographic subgroups are further fragmented into motivational and special interest segments), would compare general viewing motivations for preferred content categories to suggest what promotion strategies, program content and production qualities are most likely to attract lower SES audiences without reducing service to higher SES groups already viewing.

Conclusions

Several national uses of the new qualitative data appear highly relevant to national planning and evaluation activities. These applications await the development of a national panel, and successful quota sampling of PTV viewers and selected minority group members. The most serious stumbling block would seem to lie in the effective sampling of sufficient numbers of natural viewers of individual PTV programs. For programs with exceptionally small sizes, it may be necessary to develop a "qualitative cume" which aggregates viewers and qualitative ratings for such programs across separate program episodes. Where specials or short-lived series with only a few episodes preclude such a procedure, it may be necessary to plan for periodic recruitment of a subsample of panel members to view and evaluate such programs.

The qualitative rating scales employed in the pilot test ("entertaining," "different from most television programs," "useful to me," "informative") yielded information which speaks directly to the extent to which the public perceives PTV programs as fulfilling a number of key industry mandates. While the procedures substituted legislative and policy language for more colloquial phrases more readily applicable to television programs, these ratings can be easily translated back to summative statements intimately related to PTV programming philosophies and objectives.

The special interest group segmentation of panelists had summative value in instances where a special interest segment responded differently to a program than the total audience did. In future studies, however, the summative value of these data would be enhanced if the number and choice of interest groups were systematically derived from analysis of current PTV program content and projected programming objectives of national program development entities like CPB's Program Fund and its Educational Services Department. As Liroff noted above, the formative value of these special interest segments will depend on future validation of group membership and its relation to viewing behavior and program evaluation.

With regard to the evaluation of minority programming, analysis of the Boston data demonstrated a new capability for special assessments of targeted programs which focus primarily on the target audience, rather than on general audience reactions.

Given problems with small sample sizes for these programs under conditions of natural viewing, special procedures may be required to ensure sufficient minority sample sizes for individual targeted programs. Additionally, future studies might introduce oversampling of older persons, Asian Americans, Native Americans or other minorities whom PTV programming obligations encompass now or in the immediate future.

Finally, formative uses of the new data for broadening PTV service to various socio-economic groups would be enhanced by several new data formatting procedures. For example, tables which rank order programs by each of the four qualitative indices (and/or by some overall qualitative score derived from them) within demographic subgroups would facilitate formative analysis of a particular audience's program format, type and content preferences. Furthermore, if data analyses reveal that demographic subgroups differ significantly in their qualitative evaluations of programs, qualitative ratings for individual program should be indexed against both the total sample's average ratings and average qualitative ratings by the subgroup. Yet another valuable data formatting practice would present a merged listing of PTV and CTV qualitative indices within individual television program types (e.g., news, documentary, performance).

In the final section of this report, Dr. Joseph Philport, Arbitron project manager for the Boston field test, examines methodological issues related to the validity and efficiency of its data collection and analysis procedures. Refraining from more basic research needed to further conceptual and theoretical progress in qualitative television ratings, Philport's analysis focuses instead on those methodological concerns related to cost-effective qualitative data collection and valid applications of resulting data.

IV. METHODOLOGICAL ASSESSMENT OF FIELD TEST PROCEDURES

- Joseph C. Philport, Ph.D.
The Arbitron Company

Levels of Cooperation

Results indicate that the methodology produced a relatively high level of cooperation. Ninety percent of the households in the sampling frame of previous Arbitron in-tab sample households verbally consented to participate in the survey at the time of the diary placement telephone contact. Accordingly, diaries were forwarded to 974 persons, 12 years of age or older, who were living in consenting households during the survey period. Of those persons, 60% returned usable diaries. Although the original sampling objective was a total sample of 500 diary keepers, high cooperation rates netted a total sample of 587 viewers.

	<u>Usable Designated Sample of Households</u>	<u>Verbal Consent Rate (HH)</u>
PTV Viewing H.H.'s*	314	90% (282)
Non-PTV H.H.'s	139	89% (124)
Total	453	90% (406)

	<u>Consenting Persons</u>	<u>In-Tab Diaries</u>	<u>Return Rate</u>
PTV Viewers**	677	401	59%
Nonviewers of PTV	297	186	63%
Total	974	587	60%

*PTV Households were operationally defined as households that viewed two or more quarter-hours of public television during a previous Arbitron survey week.

**PTV Viewers were operationally defined as persons living in households that viewed two or more quarter-hours of public television during a previous Arbitron survey week.

Compared to their non-ethnic counterparts, black and Hispanic consenters were much less likely to return usable diaries. The failure to achieve in-tab objectives for Hispanics was probably the result of the use of a mono-lingual English language diary in this pilot. In future studies an improvement in Hispanic cooperation might be accomplished with the use of a special Spanish language diary. The return rate among blacks (44%) -- which was high enough to achieve the objective of 50 black panelists but somewhat less than the total sample return rate (60%) -- could be improved by special survey treatments such as additional respondent contact or special premiums.

ETHNIC RETURN RATE

	<u>Consenting Persons</u>	<u>In-Tab Diaries</u>	<u>Return Rate</u>
blacks	119	52	44%
Hispanics	57	11	19%

Effectiveness of PTV Oversampling Procedure

The sampling design included an oversampling of households that viewed two or more quarter-hours of PTV during a previous Arbitron survey. As a result, 402 persons or 68% of the in-tab sample were predesignated as PTV viewers.

In total, 44% of the sample actually viewed PTV during the week of the pilot study. Half of the predesignated PTV viewers (202 persons) and an additional 57 persons predesignated as nonviewers of PTV together composed the 259 panelists who viewed PTV during the pilot study.

DISTRIBUTION OF PERSONS WHO VIEWED PTV DURING PILOT WEEK

	<u>In-Tab</u>	<u>Persons Viewing PTV</u>	<u>Percent of Group In-Tab</u>
PTV Sample	401	202	50%
Non-PTV Sample	186	57	31%
Total	587	259	44%

While the sampling plan did generate a larger in-tab sample of PTV viewers than if a strictly random sample had been drawn, the effective sample size for PTV viewers (259) constituted only 44% of the sample, falling short of the original objective of 70% (350 PTV viewers out of 500 persons in the total sample). A more stringent definition of PTV viewing (perhaps 4 quarter-hours per week/per household) might lead to a

greater concentration of PTV viewers in the total in-tab sample, but it would undersample light viewers of PTV whose viewing and evaluation behavior might well differ from those of heavier viewers.

Non-Response on the Qualitative Scales

Respondents were instructed to evaluate each program they viewed during the survey week. Specific instructions explained how each of the four evaluative scales should be used in the evaluation of each program. (See sample diary in Appendix A.)

In total, the sample viewed 4,354 programs (duplicated), allowing for 17,416 potential qualitative responses along the four scales. Panelists completed 87% of the total number of qualitative responses possible (15,152). However, there were a number of occurrences where respondents did not use all of the scales to evaluate a program.

NON-RESPONSE IN PROGRAM EVALUATIONS

<u>Qualitative Scales</u>	<u>Number of Non-Response Occurrences</u>	<u>Percent of Total Evaluation</u>
Entertaining	371	9%
Different	627	14%
Useful	606	14%
Informative	575	13%
All Four Scales	2,179	12.5%
(Total Evaluations Possible = 17,416)		

The "entertaining" scale was used most frequently. The three remaining scales exhibited slightly higher and approximately identical levels of non-response.

Interrelationships Among Qualitative Scales

Correlation analyses were conducted to determine whether the four qualitative rating scales were measuring independent or related elements of viewer response to TV programs. The correlations* on the following page provide indications of the degree to which the qualitative scales are duplicative.

*These correlations measure the degree to which ratings along the four diary scales are interrelated. Correlations can range from -1.00 to +1.00. The larger the number, the stronger the relationship between panel responses to the two scales. A positive correlation between scales A and B means that a viewer who assigns high program ratings on scale A will tend to assign high program ratings on scale B and low ratings on scale A tend to be associated with low ratings on scale B. A negative correlation means that high ratings on scale A are associated with low ratings on Scale B, and low A ratings with high B ratings. A correlation hovering around zero indicates there is very little relationship between viewer responses on the two scales.

PEARSON PRODUCT CORRELATIONS
OF EVALUATIONS ACROSS ALL PROGRAMS

	<u>Entertaining</u>	<u>Different</u>	<u>Useful</u>	<u>Informative</u>
Entertaining	-	-	-	-
Different	.45	-	-	-
Useful	.18	.31	-	-
Informative	.12	.29	.79	-

PEARSON PRODUCT CORRELATIONS
OF EVALUATIONS ACROSS PTV PROGRAMS

	<u>Entertaining</u>	<u>Different</u>	<u>Useful</u>	<u>Informative</u>
Entertaining	-	-	-	-
Different	.40	-	-	-
Useful	.31	.32	-	-
Informative	.26	.38	.62	-

PEARSON PRODUCT CORRELATIONS
OF EVALUATION ACROSS CTV PROGRAMS

	<u>Entertaining</u>	<u>Different</u>	<u>Useful</u>	<u>Informative</u>
Entertaining	-	-	-	-
Different	.47	-	-	-
Useful	.16	.28	-	-
Informative	.09	.22	.82	-

Although there is a modest association between the "entertaining" and "different" program evaluations, the correlations suggest that it would be most appropriate to consider the two scales as separate dimensions of program evaluation.

However, the correlations between "usefulness" and "informative" scales suggest that the scales are duplicative. It is important to note that there is greater duplication between these scales in the evaluations of CTV programs. Viewers were more likely to recognize the distinction between these scales in their evaluations of PTV programs.

Qualitative Indices and Audience Size

One theoretical concern and applications-oriented question treated in this pilot test was a determination of the relationship between the size of the audience a program obtained and the magnitude of its qualitative ratings. Both researchers and programming decisionmakers have argued that qualitative scales should be more than a simple surrogate for a behavioral measure of audience size, but that they should also evidence some meaningful relationship with actual viewing behavior.

It must be noted that results presented in the following table are tentative due to small sample size. The correlations below present the degree of association between the size of the adult audience for each program and the program's subsequent qualitative ratings.

CORRELATIONS BETWEEN AUDIENCE
SIZE (ADULTS) AND QUALITATIVE PROGRAM INDICES

	CTV N=48	PTV N=43	ALL PROGRAMS N=91
Entertaining	.38*	.15	.16
Different	-.26*	-.17	-.31*
Useful	-.54*	-.14	-.44*
Informative	-.53*	.05	-.41*

*Significant at the .05 confidence level.

Comparisons of the relationships between qualitative scale ratings and size of panel audience for commercial versus public television programs reveal that the strongest associations exist between audience size and qualitative ratings for CTV programs.

In general, commercial TV programs with larger audiences were likely to be evaluated as not "informative," not "useful," not "different," but "entertaining."

However, the same correlational analysis suggests that the associations between audience size and the qualitative ratings for PTV programs are weak. Each of the qualitative indices appears to be independent of audience size; audience size for PTV programs does not appear to vary directly with viewers' qualitative reactions to the programs.

The preceding analyses examined the strength of relationship between individual qualitative scale ratings and size of program audience. Another important question about possible relationships between qualitative ratings and audience size is whether some combination of qualitative scale ratings can be used to predict the size of a program's audience. Multiple regression analyses* were conducted using the four qualitative indices as predictor variables in determining the viewing index.

As the tables below indicate, the combination of qualitative indices used in this pilot study can significantly predict the size of the audience for both CTV and PTV programs. However, the indices have different predictive values for commercial versus public television programs.

MULTIPLE REGRESSION ON SIZE OF
VIEWING INDEX -- CTV PROGRAMS

MULTIPLE R	R ²	F	df	SIGNIFICANCE LEVEL
.77	.59	15.839	4	.000

Variable	Beta
Entertaining	.82
Informative	.04
Different	-.69
Useful	-.22

Approximately 60% of the variance in CTV viewing indices (size of panel audience) could be explained by the program ratings on the four qualitative indices. CTV programs which had higher viewing indices were perceived to be high in entertainment value, not very different from most TV programs and somewhat less useful.

*Multiple regressions in this study examine the relationships among scores on the four qualitative scales and size of program audiences. Beta weights indicate the independent contribution of the four kinds of scores to accurate predictions of a program's audience size; the higher the beta weight, the more useful a scale is in predicting audience size.

MULTIPLE REGRESSION ON SIZE OF VIEWING INDEX
-- PTV PROGRAMS

MULTIPLE R	R ²	F	df	SIGNIFICANCE LEVEL
.60	.37	5.320	4	.002

<u>Variable</u>	<u>Beta</u>
Entertaining	-.07
Informative	1.49
Different	-.31
Useful	-1.21

A considerably smaller percent (37) of the variation in PTV viewing indices can be explained by the variation in the qualitative rating scale indices. Furthermore, the prediction of audience size for PTV programs is expressed by a different pattern of qualitative scores along the four diary scales. Unlike CTV programs, PTV programs with high viewing indices are perceived to be more "informative," less "useful" and, with a weaker predictive value, not very "different from most TV programs."

Motivations for Viewing Television

Information about why each respondent watched television was collected in the form of six statements of motivations for television viewing.

The principal reason these scales were incorporated was to determine if the viewing (or non-viewing) of PTV could be predicted from viewer motivations, since such possible differences between viewers and nonviewers would have direct implications for promotion of current programs, and for the production and promotion of new programs targeted to specific demographic or special interest audience segments.

Following are the results of a discriminant analysis conducted to determine whether diary keepers can be classified correctly as PTV viewers or nonviewers, based on their stated motivations for viewing television generally.

DISCRIMINANT ANALYSIS* CLASSIFYING PTV
VIEWING BY TV VIEWING MOTIVATIONS

Canonical Correlation	χ^2	df	SIGNIFICANCE LEVEL
.16	14.96	6	.02

<u>Viewing Motivations</u>	<u>Standardized Discriminant Function Coefficients</u>
to learn about people, places and things.	.79
to keep up with issues and events.	-.63
to be exposed to the arts.	.53
to give me things to talk about with other people.	-.50
to be entertained.	-.14
to pass time and relax.	-.55

NUMBER CORRECTLY PREDICTED BY
DISCRIMINANT FUNCTION (N=588)

<u>Actual Group</u>	<u>Predicted Group Membership</u>	
	<u>Non-PTV</u>	<u>PTV</u>
Non-PTV Viewers	111 59.7	75 40.3
PTV Viewers	182 45.3	220 54.7

Total Percent Predicted Accurately . . . 56.3%

*A discriminant analysis assesses the value of multiple measures in attempting to assign individuals to two or more exclusive groups -- in this case to PTV viewer and nonviewer categories.

The analysis indicates that PTV viewers were more likely to turn to TV for learning about people, places and things and for exposure to the arts. They were less likely to turn to the medium to simply pass time, to obtain topics for conversation or to keep up with issues. These data suggest that PTV viewers are more likely to use TV to expand their social and cultural perspectives than their counterparts who do not watch PTV.

While the discriminant analysis proved statistically significant, its ability to predict PTV viewing is little better than chance. It must also be noted that the cross contamination of predesignated PTV viewers and nonviewers (which occurred in the attempted sampling "prediction" of PTV viewership from previous market study data) seriously contaminated the data in this discriminant analysis. In future studies, it would be interesting to examine additional motivational items and determine whether a stronger prediction could be made among frequent, occasional, and nonviewers of PTV.

Special Interest Segments

Respondents were asked to identify the degree of their interest in 22 special interest areas. A factor analysis* was conducted to determine if this broad array of interests could be reduced to a smaller number of interest areas. The results of the factor analysis are presented in the following table.

FACTOR ANALYSIS OF SPECIAL INTERESTS

FACTOR

<u>SPECIAL INTEREST AREA</u>	<u>LOADING</u>
I. <u>CONCERN FOR PEOPLE</u>	
Children	.65
Persons with Physical and Emotional Disabilities	.72
Women's Issues	.70
Race Relations/Ethnicity	.61
Aging/Concerns of Older People	.59

*A factor analysis is used to simplify a data base by identifying a smaller number of underlying variables within the data base's total number of measures. These underlying variables, or factors, tell us which measures are describing much the same thing and the extent to which they duplicate each other. In the factor analysis above, the higher a special interest loading on a factor, the stronger its relationship to that factor.

FACTOR ANALYSIS OF SPECIAL INTERESTS
(continued)

<u>FACTOR</u>	<u>SPECIAL INTEREST AREA</u>	<u>LOADING</u>
II.	<u>WORLD AFFAIRS/POLITICS</u>	
	National News and Public Affairs	.69
	National Politics	.81
	Local Politics	.74
	International News	.66
	Local News/Community Affairs	.53
III.	<u>SCIENCE/ECOLOGY/THE ARTS</u>	
	Nature and the Environment	.63
	Arts and Culture	.53
	Science and Technology	.73
IV.	<u>FAMILY/HOUSEHOLD MANAGEMENT</u>	
	Energy	.41
	Health/Nutrition/Medical Care	.47
	Household Management	.55
	Cost of Living	.58
V.	<u>SPORTS/ATHLETICS</u>	
	Sports/Athletics	.56

The factor analysis reduced the original twenty-two special interests into five broad dimensions. The primary advantage of this procedure is that it provides a more valid definition of interest segments. However, the disadvantage is that the five dimensions represent rather broad interest areas and are therefore less functional in assessing a program's ability to reach a specific target audience. For example, if a program is targeted to the viewers interested in the Physically Disabled, it would clearly be more appropriate to segment the total audience by the appropriate special interest, rather than by the Concern for People factor.

In future studies, when a large sample size provides sufficient variation in both demographics and television program selections, it would be most useful to examine two important and related questions:

- 1) whether special group membership provides additional information beyond that which can be more simply determined from audience demographics; and
- 2) whether special interest groups, as defined by stated preferences, actually differ in their selection and evaluation of TV programs.

Recommendations for Future Testing

The pilot indicated that the oversampling of PTV viewers, based on previous viewing behavior, does lead to more PTV viewing observations than would have occurred had this control not been implemented. However, the oversampling of PTV viewers does produce a bias that weakens the validity of the qualitative indices for CTV programs and the comparisons between PTV and CTV programs. This bias can be corrected by sample weighting. Weighting would correct the distribution of PTV viewers and nonviewers in the in-tab sample while maximizing the number of observations of PTV programs.

Correct representation of Hispanics can only be obtained if a Spanish language diary is designed and incorporated in future studies.

The low return rate of diaries among black respondents could be improved by special survey treatments such as additional respondent contacts or special premiums.

The pilot results indicate that the use of four scales did not have a negative impact on scale response. The correlational analyses indicate that the only scales that are duplicative are "useful" and "informative," although subsequent regression analysis suggests the unique functions of each of these scales in predicting audience sizes for PTV programs. Reducing the number of qualitative scales would simplify the diary. However, within the limited scope of the pilot results, it appears that the four scales should be retained.

A clear identification of reporting standards on the number of viewers evaluating a program could not be identified from this study. It is therefore recommended that flagging procedures be defined for the next study, so that the user can identify qualitative program indices derived from in-tab samples of less than 70, 50, 25 and 10. The establishment of a minimal reporting standard should be a component of a future study after the measurement instrument has been refined.

The pilot results indicate that the motivational scales incorporated in the diary did not serve as strong predictors in identifying the PTV viewers. They appear to be the least functional component of the test instrument.

The pilot demonstrates that the 22 special interests included in the diary can be reduced to five general dimensions of social interest. However, the nature of the data reduction procedure would reduce the system's ability to track a program's performance within a specifically defined target audience.

It is appropriate to use both the single special interest scales and special interest factors for audience segmentation. The former provides more specific targeting information, while the latter provides more reliable broad-based data.

This final section of the report summarizes the results of the three analytic perspectives used to evaluate the field test: the local station perspective, the national perspective and methodological questions of cost-effectiveness. Capabilities and limitations demonstrated in the Boston field test are used to draw implications for future applied research in qualitative television ratings.

V. CONCLUSIONS

The original purposes of the field test described in this report were: 1) to assess the validity and efficiency of a particular qualitative data collection method, and 2) to examine the operational policy relevance of resulting data for programming executives and policymakers, at both local and national levels.

The following summary of what we learned in this exercise treats separately: a) the demonstrated capabilities of the method tested, b) limitations of those procedures, and c) implications for future research in this area.

A. Capabilities Demonstrated in the Field Test

One of the most important questions about qualitative research is whether it produced new information for programmers, producers and policymakers which would not have been available from standard program ratings information. A related, practically inseparable question is whether such new information relates to programming performance criteria unique to the public broadcasting industry. What we are really asking, in both senses, is whether the kind of qualitative data covered in the field test represents a beneficial extension of traditional program ratings data for public television. Part of the rationale for incorporating qualitative research in PTV programming and policy decisions was the belief that both industry missions and viewer behavior are more complex than traditional ratings can reveal. Are they?

We have seen that qualitative profiles for both PTV and CTV programs can differ sharply from traditional program evaluations based on audience size. Frequently, small traditional ratings were associated with high qualitative ratings, and large audience sizes occurred for programs with average or below average qualitative profiles. Additionally, viewers seemed to be able to make sense of the task of applying the scales to the programs they viewed, in the sense that all four scales had almost identical and satisfactorily high response rates.

Individually and combined, the four qualitative ratings did seem to bear a meaningful relationship to CTV viewing behavior, but were much more independent of PTV audience size. While this means the PTV programmer and policymaker are left without a simplistic "cookbook" for building audiences, these findings do add empirical credence to the philosophical roots of qualitative research--illustrating how PTV programming performance cannot and need not be crudely measured with traditional ratings information. The four new scales did produce new information.

Furthermore, the pilot test demonstrated that qualitative data can indeed speak partially to local and national accountability criteria like diversity, alternative service, education and innovation. Comparisons of CTV and PTV programs within and across similar program types, formats, dayparts and audience compositions can provide PTV with measures of alternative service in satisfying the viewing interests of various demographic and special interest segments. They also permit more fine-tuned and equitable evaluation of special interest and minority targeted programs, which cannot be evaluated with performance criteria for general audience programs. While certain methodological problems must be overcome before the data are completely responsive, the pilot test established the feasibility of empirical evaluation of some of the long-needed PTV industry performance standards ignored by traditional ratings research.

Fine-tuning of research data to PTV accountability criteria is also furthered by the special interest audience segmentation feature of the diary. This psychographic segmentation allows new, non-demographic approaches to studying PTV programming diversity.

Finally, in addition to such summative applications, we have seen formative applications of the new data. While the ratings scales, special interest and viewing motivation data were intended to strike a middle ground between traditional ratings and more extensive measures of program impact, they do seem useful in "flagging" potential problems with a program which may require more extensive and specialized study in order to improve its performance (as in Liroff's reactions to the Ten O'Clock News program). Also, with improved sampling procedures, it is feasible to use the new data to replace stereotypical notions about different types of viewers and to recommend program formats, content, scheduling and promotional messages most likely to initiate or improve service to a specific target audience.

B. Limitations of the Methodology and Data Reporting Practices

Sample Size for Individual PTV Programs

The largest and most pervasive problem with the data was the small size of the panel sample. While the sample size was determined to satisfy the methodological issues brought to the pilot test and not to make representative statements about program evaluation, serious problems arose when analyses dealt with segments of the total sample.

The most serious problem was the small sample sizes obtained for individual PTV programs. The procedure for oversampling PTV viewers did produce a larger number of PTV viewers than would have been obtained with random sampling, but it fell short of the original number of PTV viewers desired. This problem at least partially contributed to the chronically insufficient sample sizes found for individual PTV programs.

Small sample sizes for individual PTV programs were also a function of the natural viewing behavior encouraged from the sample and the one-wave nature of the field test. The sample size problem is magnified, of course, in data analysis which further segments PTV viewers of a given program by demographics or special interest membership or viewing motivation. For instance, the lower cooperation rate of black and Hispanic viewers coupled with their lower frequency of PTV viewing made analyses of their responses to public television impossible. While a large national sample could obtain more minority response to PTV programs through oversampling procedures, it is apparent that both incidence in the general population and frequency and amount of PTV viewing must be taken into account when developing quota samples and special recruitment procedures for any demographic subgroups of interest. Furthermore, minimum reporting standards should be established for qualitative program ratings, below which sample size requires very cautious interpretation of the data.

Minority Sampling Problems

Generally, cooperation for the total sample was satisfactorily high. However, Hispanic cooperation was notably lower, probably due to the lack of a bilingual diary or special recruitment procedures. Black cooperation was satisfactory, but somewhat lower than that of the total sample. Given less frequent PTV viewing by black panelists, special recruitment to view individual programs or a special quota sampling of black PTV viewers (and nonviewers) might be required.

PTV Viewers/Nonviewer Sampling Problems

Considerable cross-contamination of predicted viewers and nonviewers of PTV prevented a variety of interesting analyses (e.g., possibly different evaluation standards for CTV, different TV viewing motivations, different preferred program types or formats). If a procedure could be found to discriminate between, and proportionately sample PTV viewers and nonviewers, one important formative application would involve audience building. At both local and national levels, audience expansion efforts could be facilitated with comparisons of the viewing motivations, program preferences and areas of special interest among PTV viewers and nonviewers. A related and intriguing question (related to Liroff's analyses of This Old House) is how programming and promotion activities can be designed to attract the nonviewer, without alienating, or decreasing service to the current viewership.

Finally, when valid procedures have been determined, weighting calculations will be required to bring the total sample, with its overrepresentation of PTV viewers (and minority viewers), into alignment with the total market or national population to which findings are being generalized.

Reporting Formats

A variety of improvements can be made in data formatting to enhance the applied value of the research data. The notation of certain sample sizes as "below measurement standards" has already been mentioned. A related improvement would be to indicate exact sample size for each program next to the viewing index, so that users can quickly determine the number of persons contributing to each program's qualitative rating.

Third, given both the formative and summative applications of the data attempted in Section II and III, it would be helpful to group the same data in additional patterns. One such improvement would be a graphic representation of all program ratings for each of the four qualitative scales. Another would report merged data for CTV and PTV programs of the same type, including average type ratings, so that program types could be judged against standards particular to that genre. A related improvement would entail the derivation of one general qualitative number which, in a statistically equitable fashion, provides an overall qualitative rating for each program. The full four-element qualitative scores would, of course, still be reported for those users who prefer more detailed information on a given program.

Finally, given further improvement of the special interest group data (as described below), merged CTV and PTV qualitative data could be studied to evaluate existing service to that interest group and to determine areas of superior and inferior service. Programs with content known to be related to the special interest topic should be highlighted.

Validation Procedures

External validity of qualitative ratings occasionally puts the researcher in a double bind. On the one hand, if there is a strong relationship with audience size, parsimony and cost-effectiveness speak against the collection of any new data. On the other hand, if there is no relationship between program evaluations and program viewing behavior, critics rightly question the practical worth of the new data and occasionally dismiss it as meaningless verbal artifact. A fresh look at the issue would predict a relationship complicated by a host of intervening variables (e.g., competing nonviewing activities, noncontrol of program choice, counterprogramming), such that well liked programs are not necessarily highly viewed, and highly viewed programs are not necessarily well liked. An eyeballing of these study data certainly uncovered these various possibilities, and more sophisticated statistical analyses revealed a complex but meaningful relationship between viewing levels and qualitative ratings (pp. 36-37).

In terms of internal validity considerations, the one-wave nature of the field test did not permit measurements of the reliability of diary keepers' responses. Especially considering the possibility of inter-episode variation in a series' concepts and treatments, it will be essential to plan repeated qualitative evaluations of a series. Given the high incidence of PTV series with a small number of episodes, this is a planning consideration which recommends repeated survey waves and possibly special procedures for short-lived series.

Additional needs for validity information arose frequently in relation to the special interest group feature of the diary, where, for both formative and summative purposes, users of the data will require certainties about the relationship between special interest group membership and actual viewing and evaluation behaviors. While these scientific assessments of this diary feature await further study, its local and national policy relevance is clearly acknowledged.

Yet another validity concern arises in relation to general TV viewing motivations. Given the methodological problems in distinguishing viewers and nonviewers of PTV, it is difficult to assess the validity or utility of this diary feature. Further empirical analysis of possible relationships between viewing motivations and viewer demographics are also required.

A related validity concern uncovered in Philport's analysis was the similarity in respondents' patterns of use for the "informative" and "useful" scales. The apparent redundancy of the two qualitative criteria suggests that a new and less duplicative scale might be substituted in future studies in order to maximize the audience information obtained, unless the summative policy relevance of both measures recommends no alterations.

C. Implications for Future Research

Problems of insufficient sample size for specific PTV programs might be approached in several ways. One approach might be to aggregate data from repeated survey waves to establish a "qualitative cumé" for a program. This approach would work for programs aired long enough to permit repeated measurements and it would help account for inter-episode variations in program content and production quality.

However, for short-lived series, specials or programs with especially low viewing levels, special practices may be needed to capture a sufficient sample of viewers. Especially with targeted programs, it may be helpful to imitate the qualitative ratings procedures of our European counterparts, who use special segments of the diary for in-depth, diagnostic questions on one or two programs of special interest during a given survey wave.

Another sample problem in the pilot test concerned the successful differentiation of PTV viewers and nonviewers. Philport pointed out that more stringent definitions of PTV viewers (e.g., by requiring larger amounts of PTV viewing per month or week) might underrepresent light PTV viewers' opinions. Another approach might use a special PTV viewing instrument administered to PTV viewers predesignated with earlier diaries, or the definition of nonviewers might use two or three previous diary waves to define nonviewing over a longer period of time.

A final sampling issue deals with programs. While the current test allowed for evaluations of the output of seven television outlets, as new technologies increasingly penetrate the telecommunications environment, the system will have to accommodate program evaluations for software from cable, videocassettes and video-discs, with all the program coding and analysis issues that result.

Appendix A

Qualitative Diary Excerpts

Diary format for television viewing motivations
and special interest group membership

Please complete the following questions so that broadcasters can know more about the interests of their audiences.

1 Please indicate how often each one of the following reasons applies to why you watch television.

- To learn about people, places and things
- To keep up with current issues and events
- To be exposed to the arts
- To give me things to talk about with other people
- To be entertained
- To relax and pass the time

Never Applies	Occasionally Applies	Generally Applies	Always Applies
☐	☐	☐	☐
☐	☐	☐	☐
☐	☐	☐	☐
☐	☐	☐	☐
☐	☐	☐	☐
☐	☐	☐	☐

2 Please indicate how interested you are in each of the following subject areas.

	Not at all Interested	Somewhat Interested	Very Interested	Extremely Interested
Energy	☐	☐	☐	☐
Children	☐	☐	☐	☐
Crime and Public Safety	☐	☐	☐	☐
National News and Public Affairs	☐	☐	☐	☐
Religion/Ethics/Morality	☐	☐	☐	☐
Health/Nutrition/Medical Care	☐	☐	☐	☐
Nature and the Environment	☐	☐	☐	☐
National Politics	☐	☐	☐	☐
Local Politics	☐	☐	☐	☐
History/Biography	☐	☐	☐	☐
International News	☐	☐	☐	☐
Education and Training	☐	☐	☐	☐
Local News/Community Affairs	☐	☐	☐	☐
Concerns of Persons with Physical or Emotional Disabilities	☐	☐	☐	☐
Women's Issues	☐	☐	☐	☐
Arts and Cultural Activities	☐	☐	☐	☐
Science and Technology	☐	☐	☐	☐
Household Management	☐	☐	☐	☐
Sports/Athletics	☐	☐	☐	☐
Cost of Living/the Economy	☐	☐	☐	☐
Race Relations/Ethnicity	☐	☐	☐	☐
Aging/Concerns of Older People	☐	☐	☐	☐
Other (specify)	☐	☐	☐	☐

Please answer the questions on the next page...

Diary format for one day's evaluation of viewed
programs along four qualitative rating scales

Wednesday
November 18, 1968

Check if viewed	Time	Channel Number	Station Call Letters	Program Title	How entertaining was each television program that you viewed?				How different from other television pro- grams was each program that you viewed?				How useful to you was each television program that you viewed?				How informative was each television program that you viewed?			
					1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
					Not at all Entertaining	Somewhat Entertaining	Very Entertaining	Extremely Entertaining	Not at all Different	Somewhat Different	Very Different	Extremely Different	Not at all Useful	Somewhat Useful	Very Useful	Extremely Useful	Not at all Informative	Somewhat Informative	Very Informative	Extremely Informative
☐	1:00-1:29PM				☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐		
☐	1:30-1:59PM				☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐		
☐	4:00-4:29PM				☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐		
☐	4:30-4:59PM				☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐		
☐	5:00-5:29PM				☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐		
☐	5:30-5:59PM				☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐		
☐	6:00-6:29PM				☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐		
☐	6:30-6:59PM				☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐		
☐	7:00-7:29PM				☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐		
☐	7:30-7:59PM				☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐		
☐	8:00-8:29PM				☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐		
☐	8:30-8:59PM				☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐		
☐	9:00-9:29PM				☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐		
☐	9:30-9:59PM				☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐		
☐	10:00-10:29PM				☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐		
☐	10:30-10:59PM				☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐		
☐	11:00-11:29PM				☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐		
☐	11:30-MIDNIGHT				☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐		

Appendix B

BIBLIOGRAPHY

- British Broadcasting Corporation. Annual Review of BBC Audience Research Findings. London: BBC Audience Research Department, 1977.
- Corporation for Public Broadcasting. Comparison of Public and Commercial Television Program Types. Unpublished Internal Memorandum, Office of Communication Research, 1979.
- Corporation for Public Broadcasting. Proceedings of the 1980 Applications Conference on Qualitative Television Ratings. Washington, D.C.: CPB Office of Communication Research, May 1980.
- Corporation for Public Broadcasting. Proceedings of 1980 Technical Conference on Qualitative Television Ratings. Washington, D.C.: CPB Office of Communication Research, March 1980.
- Corporation for Public Broadcasting. Review of 1980 CPB Communication Research Findings. Washington, D.C.: CPB Office of Communication Research, March 1981.
- Durand, J. and J. Arachtingi. "The History of Qualitative Television Ratings in France," in OCR News Brief, Vol. IV, No. 4. Washington, D.C.: CPB, April 1, 1981.
- Frank, Ronald E. and M. G. Greenberg. The Public's Use of Television. Beverly Hills: Sage Publications, 1980.
- Goodhardt, G. J., Ehrenberg, A.S.C. and M.A. Collins. The Television Audience: Patterns of Viewing. Lexington, Massachusetts: Lexington Books, 1975.
- Himmelweit, Hilde T., Swift, Betty and Marianne E. Jaeger, "The Audience as Critic: A Conceptual Analysis of Television Entertainment," in Tannenbaum, Percy H. (ed). The Entertainment Functions of Television. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1980, pp. 67-106.
- Independent Broadcasting Authority. Audience Appreciation Report. London: IBA Audience Research Department, February 1981.
- Landis, Jack. Mapping Analysis for the Corporation for Public Broadcasting. West Palm Beach, FL: The Landis Group, October 1979.
- Owen, Bruce., Beebe, J. H. and W. G. Manning, Jr. "Theories of Program Choice," in Television Economics. Lexington, Massachusetts: Lexington Books, 1974, pp. 49-88.
- Webster, James. Factors Mediating Program Preference As A Determinant of TV Viewing Behavior. Unpublished Ph.D. Dissertation. Bloomington, Indiana: Indiana University, 1980.
- WGBH Community Advisory Board. Report to the Trustees of the WGBH Educational Foundation from the WGBH Community Advisory Board. Boston, Mass.: WGBH, April 1980.